Epidemiol. Infect. (2010), **138**, 992–1003. © Cambridge University Press 2009
doi:10.1017/S0950268809991075

Binary genomotyping using lipooligosaccharide biosynthesis genes distinguishes between *Campylobacter jejuni* isolates within poultry-associated multilocus sequence types

G. S. HOTTER[1], I. H. LI[2] AND N. P. FRENCH[2]*

[1] *AgResearch, Hopkirk Research Institute, Grasslands Research Centre, Palmerston North, New Zealand*
[2] *Institute of Veterinary, Animal and Biomedical Sciences, Hopkirk Research Institute, Massey University, Palmerston North, New Zealand*

(*Accepted 7 October 2009; first published online 3 November 2009*)

SUMMARY

Campylobacter jejuni is a leading cause of human bacterial gastroenteritis throughout the industrialized world. We investigated whether or not differences in gene complement at the lipooligosaccharide (LOS) biosynthesis locus can identify epidemiologically useful binary genomotypes in 87 *C. jejuni* isolates from poultry-associated multilocus sequence types (STs) collected during the course of a sentinel surveillance study. Using a PCR-based approach, we correlated assignment of both isolate LOS locus class and binary genomotype with ST. We found that isolates within STs 45, 190, 354 and 474 displayed mosaicism in gene complement at the intra-ST level. For example, based upon their binary genomotypes, we assigned individual ST-45 isolates from human clinical cases as probably originating from either a poultry or wild-bird source. However, intra-ST mosaicism in gene complement was observed alongside broader patterns of congruence in LOS locus class and gene complement that distinguished between isolates from different STs.

Key words: *Campylobacter*, epidemiology, foodborne zoonoses, public health microbiology, zoonoses.

INTRODUCTION

Campylobacter jejuni is a leading cause of human bacterial gastroenteritis throughout the industrialized world. In New Zealand, the campylobacteriosis national notification rate was >300 people/100 000 population for the 5 years preceding 2008, with peaks within areas covered by individual district health boards of >500 people/100 000 population [1, 2]. The year 2008 saw a significant decrease in the notification rate coincident with interventions within the poultry industry [3, 4]. *C. jejuni* is also associated with the development of post-infection complications such as reactive arthritis and Guillain–Barré syndrome (GBS), GBS being the most frequent cause of general paralysis since the near-worldwide eradication of poliomyelitis [5, 6] and being observed in about two people per 100 000 population in New Zealand each year [7]. Structural mimicry between sialylated *C. jejuni* lipooligosaccharide (LOS) and human gangliosides is thought to be one of the causes of GBS whereby LOS induces the production of autoantibodies against the structurally related gangliosides [5, 6]. In addition to the involvement of LOS in GBS, LOS is also implicated in virulence, including invasion of epithelial cells [8, 9]. Furthermore, the LOS biosynthesis locus is one of several genomic

* Author for correspondence: Dr N. P. French, Professor of Food Safety and Veterinary Public Health EpiCentre, Hopkirk Institute, Institute of Veterinary, Animal and Biomedical Sciences, College of Sciences, Massey University, New Zealand.
(Email: N.P.French@massey.ac.nz)

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

hotspots identified for variation in gene complement between *Campylobacter* spp. and between *C. jejuni* strains [10–19].

Molecular epidemiological investigations of *C. jejuni* have utilized a range of DNA sequence-based methods to characterize isolates including: PCR–RFLP based upon *flaA* (flagellin gene), pulsed-field gel electrophoresis (PFGE), amplified fragment length polymorphism (AFLP), and multilocus sequence typing (MLST) based upon sequencing 400–500 bp regions within seven housekeeping genes (loci) [20, 21]. MLST has some significant advantages, such as inter-laboratory reproducibility and portability of results, which has permitted establishment of the internationally accessible PubMLST database. MLST is particularly effective when applied to longitudinal studies of population structure. Furthermore, recent studies performed in England, New Zealand, and Scotland have shown that comparison between the sequence types (STs) assigned to human clinical isolates and the STs assigned to environmental and animal isolates allows MLST to be used to assign human isolates to source populations on a probabilistic basis [22–27]. Indeed, many of the most prominent multilocus STs, such as STs 42, 45, 48, 61 and 354, isolated from human clinical cases and from animal samples in New Zealand are also among the most commonly observed in other countries [22, 24, 28]. However, some STs that are prominent in New Zealand are comparatively rare elsewhere including, in particular, ST-474, the most common single ST associated with human clinical campylobacteriosis in New Zealand [22, 24, 28]. For example, a sentinel surveillance study conducted from 2005 to 2008 in the Manawatu District of the lower North Island of New Zealand applied source attribution modelling methods to attribute some 70–80% of human campylobacteriosis cases to a poultry source [22–24]. Moreover, within that study, STs such as ST-474, ST-354 and ST-45 were classified as poultry-associated, and STs such as ST-190 were classified as both ruminant- and poultry-associated [22–24], and these four STs form the primary focus of the current study.

While recent application of source attribution models to MLST data has facilitated the assignment of human isolates to source populations on a probabilistic basis, application of MLST may be limiting in situations where more detailed population structure is required, as might be the case when seeking to identify the specific source of an outbreak. In such cases MLST can be supplemented by the analysis of additional, more variable sequences such as the *flaA* and *flaB* short variable regions (*flaA* SVR; *flaB* SVR) and *porA* [29]. While application of MLST and extended MLST methodologies are beginning to inform us about the relative contributions of different transmission routes through animal food chains, identification of specific sources remains problematic.

Binary genomotyping describes the use of experimental approaches to type isolates on the basis of their comparative gene complement. A number of different approaches have been used to derive binary genomotypes for *C. jejuni* isolates: genomic DNA hybridization to microarrays for genome-wide analysis of gene complement; gene-targeted PCR-based screens; or combinations of both approaches [10, 17, 30–32]. Similar approaches have been applied to other bacterial species such as *Streptococcus pneumoniae* [33] and *Escherichia coli* [34]. In the current study we adopted a PCR-based approach to binary genomotyping, assessing LOS gene complement by scoring the full complement of 42 LOS genes/open reading frames (orfs) described by Parker *et al.* [17] as either PCR-positive or PCR-negative. We applied this method to isolates from important poultry-associated STs identified in human clinical campylobacteriosis cases during the course of a sentinel surveillance study conducted from 2005 to 2008 in the Manawatu District of the lower North Island of New Zealand [22–24]. We sought to determine: if *C. jejuni* isolates indistinguishable on the basis of MLST can be distinguished on the basis of variation in their LOS gene complement; and, if so, whether the observed differences in binary genomotypes convey epidemiologically useful information.

METHODS

C. jejuni isolate selection, culture, DNA preparation, and MLST

This study focused on 87 isolates collected primarily during the course of a sentinel surveillance study conducted from 2005 to 2008 in the Manawatu District of the lower North Island of New Zealand [22–24]. The present study focused primarily on isolates from four prominent poultry-associated STs identified in that study: 35 isolates from ST-474; 12 isolates from ST-354; 17 isolates from ST-45, which is also associated with wild birds; 12 isolates from ST-190, a member of the internationally

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

Table 1. *Relationships between sequence type* (*ST*), *clonal complex* (*CC*) *and multilocus sequence typing housekeeping gene alleles for STs examined in this study*

ST	CC	*aspA*	*glnA*	*gltA*	*glyA*	*glmM* (*pgm*)	*tkt*	*atpA* (*uncA*)
354	354	8	10	2	2	11	12	6
38	48	2	4	2	2	6	1	5
3609	48	2	4	12	2	7	1	5
48	48	2	4	1	2	7	1	5
474	48	2	4	1	2	2	1	5
2350	48	2	4	1	2	2	5	5
3718	48	2	4	1	4	1	1	5
190	21	2	1	5	3	2	3	5
45	45	4	7	10	4	1	7	1

important clonal complex (CC) 21, that is responsible for a significant proportion of human campylobacteriosis cases worldwide (a CC includes a founding ST and all of its single, double and triple locus variants); and an additional 11 isolates from other STs within CC-48. Isolates were randomly selected for the current study from a larger bank of frozen isolates collected during the sentinel surveillance study and comprise totals of 45 human isolates (H or no prefix in the figure and tables), 33 poultry isolates (P or T prefix), four isolates from non-poultry meat (M prefix), four isolates from wild-bird faeces collected in recreational areas (R prefix), and one isolate from a natural waterway (W prefix). As percentages of the total number of isolates stored in the bank of frozen isolates, the ST-474, ST-190 and ST-3809 isolates we examined represent ~15% of stored isolates from those STs, the ST-354 and ST-2350 isolates we examined represent ~33% of stored isolates from those STs, the ST-45 isolates we examined represent ~10% of the stored ST-45 isolates, the ST-38 and ST-48 isolates we examined represent ~5% of stored isolates from those STs, and the ST-3718 isolate we examined is a singleton. Isolate information is provided in Tables 1 and 2, and in Supplementary Table S1 (available in the online version of the paper). Colonies, representing a third round of single colony purification from an original sample, were cultured on blood agar (Fort Richard, NZ) by microaerobic incubation at 42 °C for 2 days. Chromosomal DNA was prepared for each isolate by boiling a loop of cells for 10 min in 500 μl sterile MilliQ water. The sample was centrifuged to remove cell debris, and the supernatant was transferred to a fresh tube to use for binary genomotyping. MLST had already been performed on each isolate, as previously described [20], as part of the sentinel surveillance study [22–24]. The allelic profiles for the STs we examined together with the relationships between the STs are shown in Table 1. The ST for each isolate was assigned with the aid of the *Campylobacter* PubMLST Database (http://pubmlst.org/campylobacter/).

Binary genomotyping

PCR reactions for binary genomotyping were performed using a set of 42 previously described LOS gene primers that amplify genes/orfs (henceforth, orfs) in the region *waaC-waaF* [17]. PCR reactions were performed on a Bio-Rad iCycler using 3 μM final concentration for each primer in 25 μl reaction mixes with iTaq (Bio-Rad, USA) using the PCR cycle: 94 °C for 3 min, 35× (94 °C for 30 s, 55 °C for 30 s, 72 °C for 60 s), and 72 °C for 5 min. PCR reaction products were examined by agarose gel electrophoresis and scored as PCR-positive or PCR-negative (orf absent or divergent). Orf2 (*waaM*) and orf12 (*waaV*) served as positive controls. A PCR-negative result for an orf in an isolate could mean that either the orf is absent, or that the sequence of the targeted gene/orf in that *C. jejuni* isolate is sufficiently divergent to one or both primers used such that amplification does not occur.

Data analysis

LOS locus classes were formally assigned using a subset of 16 of the 42 primer pairs used for binary genomotyping [17]. These 16 primer pairs were: a primer pair amplifying orf12 (*waaV*) as a positive control; 12 primer pairs (orf6ab1, orf6ab2, orf7ab, orf8ab, orf5bII, orf6c, orf7c, orf8c, orf17d, orf18df, orf3e and orf27e) that permit assignment of LOS locus classes A1, A2, B1, B2, C, D and E; and three primer pairs (orf19df, orf16df and orf26e) that permit assignment of LOS locus classes F and H [17]. These primers are highlighted in bold type in Supplementary Table S1 (available online), and the LOS locus classification system based upon the PCR positive and PCR negative patterns for the 16 orfs is shown in Table 3 [17].

The relationship between isolates, based on their relative gene complement, scored as PCR-positive or PCR-negative for each of the 42 orfs (their

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

Table 2. *Lipooligosaccharide locus classes and binary genomotypes assigned to isolates from poultry-associated sequence types (STs)*

Isolate*	ST	CC	LLC	BG	Isolate	ST	CC	LLC	BG
H231	354	354	U	16	P238b-C2	474	48	B2E	4
H238	354	354	U	16	P249b-C1	474	48	B2	6
H574	354	354	U	16	P258c-C1	474	48	B2	9
H603	354	354	U	15	P259d-C1	474	48	B2	9
H626	354	354	U	17	P282b-C2	474	48	B2	9
H667	354	354	D	19	P328b-C1	474	48	B2	6
P303a-C1	354	354	E	25	P489a-C1	474	48	B2	6
P414a-C1	354	354	U	16	P554b-C1	474	48	B2	6
T15-C1	354	354	E	21	W128b	474	48	B2	9
T16-C1	354	354	U	18	M123a	474	48	B2	9
T272-C1	354	354	U	15	M727a	474	48	B2	9
T551-C1	354	354	U	16	M1039b	474	48	B2	9
H220	38	48	B2	9	H304	2350	48	B2	9
P106a-C2	3609	48	B2	9	H207a	3718	48	B2	9
P135a-C2	3609	48	B2	9	H12986a	190	21	CE	35
P145a-C2	3609	48	B2	9	H12172	190	21	CE	34
P157a-C2	3609	48	B2	9	H75	190	21	C	30
H17601a	48	48	B2	9	H127	190	21	C	30
H16956	48	48	B2	9	H787	190	21	C	31
P136a-C2	48	48	B2	9	H788	190	21	C	27
P140a-C2	48	48	B2	9	H791	190	21	CE	34
H73020	474	48	B2	13	M1272a	190	21	C	32
H22082	474	48	B2E	3	P165a-C1	190	21	C	33
H31262	474	48	B2	12	P164a-C1	190	21	C	28
H71	474	48	B2	14	P165b-C1	190	21	C	29
H74	474	48	B2	7	P427a-C1	190	21	CE	34
H95	474	48	B2	7	H75250a	45	45	E	21
H80	474	48	B2	9	H76800	45	45	E	21
H123	474	48	B2E	3	H73687	45	45	H	22
H128	474	48	B2E	3	051206734	45	45	E	23
H133	474	48	B2E	3	H272	45	45	E	24
H121	474	48	B2	10	H288	45	45	EB2	5
H143	474	48	B2	9	P10b-C1b	45	45	H	26
H169	474	48	B2	9	P11b-C2	45	45	H	26
H164	474	48	B2	13	P115b-C3	45	45	U	20
H168	474	48	B2	9	P137c-C3	45	45	H	22
H319	474	48	B2	8	P133a-C3	45	45	H	22
H333	474	48	B2	9	P135b-C2	45	45	H	22
H350	474	48	B2	9	P155b-C2	45	45	H	22
H360	474	48	B2	9	R53B	45	45	E	23
H492	474	48	B2	11	R48D	45	45	E	23
H676	474	48	B2	11	R3	45	45	E	23
P110b-C1	474	48	B2E	1	R5a	45	45	E	23
P170c-C1	474	48	B2E	2					

CC, Clonal complex; LLC, lipooligosaccharide locus class; BG, binary genomotype.

* Isolate designations: H or no prefix, isolates sourced from human clinical cases; P or T prefix, isolates sourced from retail poultry; M prefix, isolates sourced from retail beef or lamb; R prefix, isolates sourced from wild-bird faeces collected in recreational areas; W prefix, isolate sourced from an environmental waterway; C1–C3, poultry companies 1–3.

binary genomotypes), was examined by calculating the simple matching coefficient, and constructing a tree using the unweighted pair-group method with arithmetic means (UPGMA) [35]. The cluster analysis was performed using BioNumerics (Applied Maths Inc., Belgium).

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

Table 3. *Assignment of lipooligosaccharide locus classes based upon PCR-assessed orf complement*

LOS locus class	orf12	orf6ab1	orf6ab2	orf7ab	orf8ab	orf5bII	orf6c	orf7c	orf8c	orf17d	orf18df	orf3e	orf27e	orf19df	orf16df	orf26e
A1	+	+	−	+	+	−	−	−	−	−	−	−	−	−	−	−
A2	+	−	+	+	+	−	−	−	−	−	−	−	−	−	−	−
B1	+	+	−	+	+	+	−	−	−	−	−	−	−	−	−	−
B2	+	−	+	+	+	+	−	−	−	−	−	−	−	−	−	−
C	+	−	−	−	−	−	+	+	+	−	−	−	−	−	−	−
D	+	−	−	−	−	−	−	−	−	+	+	−	−	+	+	−
E	+	−	−	−	+/−	−	−	−	−	−	−	+	+	−	−	+
F	+	−	−	−	−	−	−	−	−	−	+	−	−	+	+	−
H	+	−	−	−	−	−	−	−	−	−	−	+	+	−	−	−

RESULTS

Assignment of LOS locus classes

We used amplification of orf12 (*waaV*, positive control) in conjunction with amplification of the 16 orfs highlighted in boldface in Supplementary Table S1, representing orfs from LOS locus classes A1, A2, B1, B2, C, D, E, F and H [17], to assign LOS classes to 87 New Zealand *C. jejuni* isolates sourced from environmental waters, retail poultry, retail beef or lamb, wild-bird faeces collected from recreational areas, and from human clinical campylobacteriosis cases (Table 2, Supplementary Table S1). The MLST alleles corresponding to the STs examined in this study are shown in Table 1.

All non-ST-474 isolates we examined from CC-48, and most ST-474 isolates, were assigned to LOS locus class B2 (Fig. 1, Table 2, Supplementary Table S1). However, seven ST-474 isolates were assigned to LOS locus class B2E to show that those seven isolates were also PCR-positive for a group of LOS locus class E orfs while retaining their LOS locus class B orfs. We examined retail poultry isolates from two poultry companies and assigned LOS locus classes B2 and B2E to isolates from both companies.

All ST/CC-45 isolates we examined from wild-bird faeces were assigned to LOS locus class E. We examined seven ST/CC-45 isolates from three poultry companies, with six of seven isolates from companies 1–3 being assigned to LOS locus class H, while the remaining isolate from company 3 could not be assigned to a LOS locus class (Fig. 1, Table 2, Supplementary Table S1). Four of six ST-45 isolates we examined from human clinical cases were assigned to LOS locus class E (H75250a, H76800, H051206734 and H272), one was assigned to LOS locus class H (H73687), and one was assigned to LOS locus class EB2 (H288) to show that it was also PCR-positive for a group of LOS locus class B2 orfs.

All ST-190/CC-21 isolates we examined were assigned to either LOS locus class C or to a LOS locus class C variant that we designated CE to show that those isolates were also PCR-positive for additional orfs, primarily from LOS locus class E (Fig. 1, Table 2, Supplementary Table S1). We observed both LOS locus classes C and CE in poultry and human isolates.

We were able to assign two poultry isolates from ST/CC-354 to LOS locus class E (P303 and T15), and an isolate from a human clinical case to LOS locus class D (H667). While the remaining nine ST/CC-354 isolates we examined could not be assigned to a specific LOS locus class [17], with orf17d being PCR-negative (Fig. 1, Supplementary Table S1), these remaining isolates are most closely related to LOS locus class F.

Binary genomotyping with LOS genes

While assignment of LOS locus classes alone captured broad patterns of relative gene complement for isolates within CC-48, ST/CC-45 and ST-190/CC-21, it did not capture all information available across all 42 orfs we examined. Furthermore, only three of 12 ST/CC-354 isolates we examined could be assigned to a formal LOS locus class.

The full PCR-positive or PCR-negative results (gene complement) for each orf for all isolates examined, together with each isolates assigned numerical binary genomotype, are shown in Figure 1 and in Supplementary Table S1. Figure 1 depicts a dendrogram clustering all 87 isolates on the basis of their

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

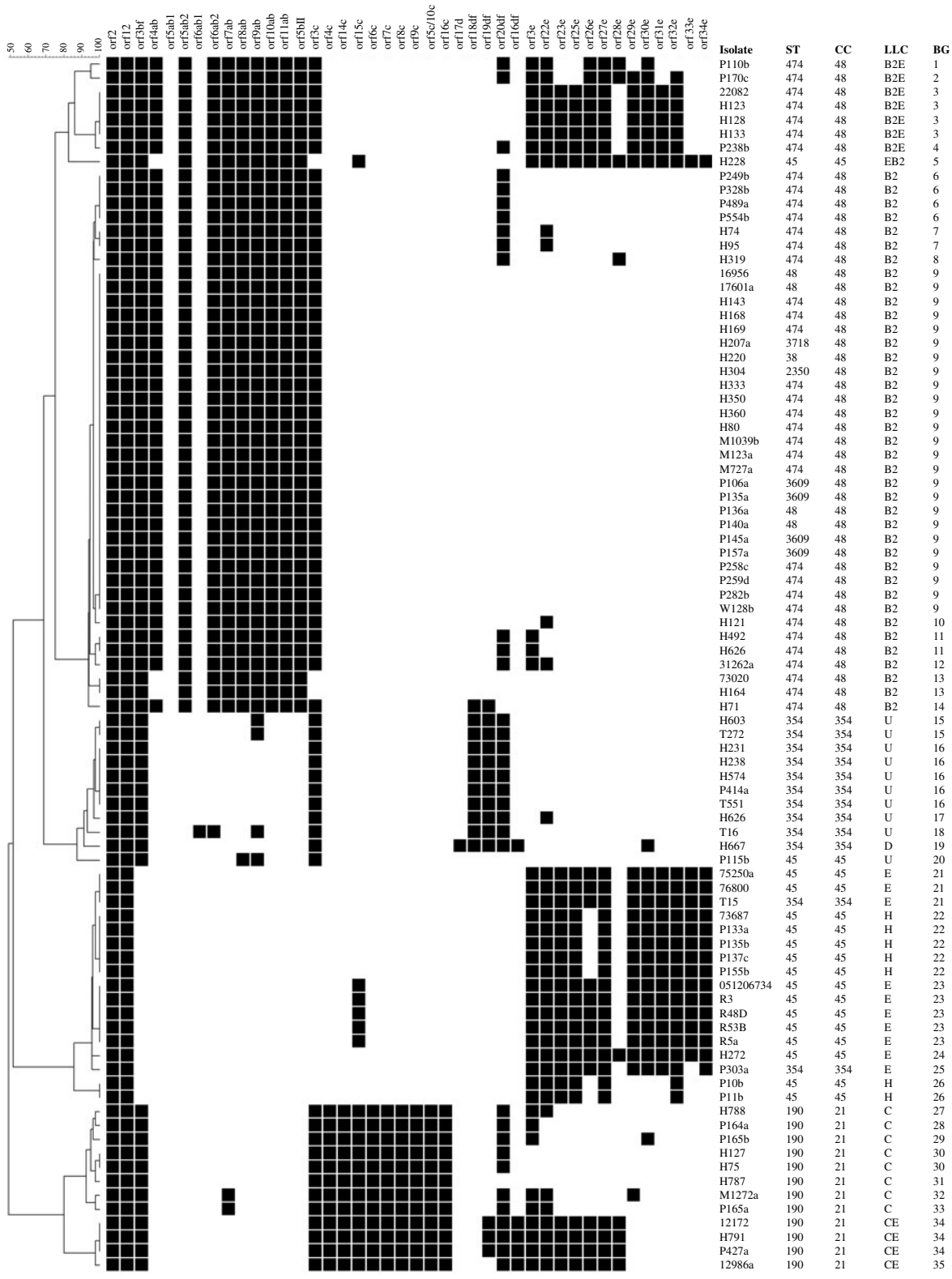

Fig. 1. Dendrogram generated using simple matching UPGMA clustering all isolates on the basis of their LOS orf complement. Black box, PCR-positive for that orf; white box, PCR-negative for that orf. H or no prefix within isolate designations, isolates sourced from human clinical cases; P or T prefix, isolates sourced from retail poultry; R prefix, isolates sourced from wild-bird faeces in recreational areas; M prefix, isolates sourced from cattle or sheep meat; W prefix, isolates sourced from environmental water sources. LLC, Lipooligosaccharide locus class; BG, numerical binary genomotype.

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

Table 4. *Summary data showing the relationships between sequence type (ST), clonal complex (CC), lipooligosaccharide (LOS) locus class and binary genomotype for isolates examined in this study*

ST	CC	LOS locus class	No. of isolates	Source(s)*	Binary genomotype(s)
354	354	U (F-related)	9	H, P, T	15, 16, 17, 18
		D	1	H	19
		E	2	P, T	21, 25
38	48	B2	1	H	9
3609	48	B2	4	P	9
48	48	B2	4	H, P	9
474	48	B2	28	H, P, W, M	6, 7, 8, 9, 10, 11, 12, 13, 14
		B2E	7	H, P	1, 2, 3, 4
2350	48	B2	1	H	9
3718	48	B2	1	H	9
190	21	C	8	H, P	27, 28, 29, 30, 31, 32, 33
		CE	4	H, P	34, 35
45	45	E	8	H, R	21, 23, 24
		EB2	1	H	5
		H	7	H, P	22, 26
		U	1	P	20

* Source abbreviations: H, isolates sourced from human clinical cases; P or T, isolates sourced from retail poultry; M, isolates sourced from retail beef or lamb; R, isolates sourced from wild bird faeces collected in recreational areas; W, isolate sourced from an environmental waterway.

gene complement, and a numerical binary genomotype was assigned to each unique gene complement according to the first record of each unique gene complement in Figure 1. Table 4 summarizes the relationships observed between ST, CC, LOS locus class, source, and binary genomotype.

All isolates were PCR-positive for the two control orfs, orf 2 (*waaM*) and orf12 (*waaV*). The dominant feature displayed by the dendrogram for the set of isolates we examined is that, with the exception of two isolates from ST/CC-45 (H288 and P115b) and two isolates from ST/CC-354 (T15 and P303a), all other isolates cluster to at least the ST/CC level, with isolates from each ST/CC having related binary genomotype(s). The four exceptions show patterns of gene loss and/or gain that result in alternative cluster relationships. Thus, ST/CC-45 isolate H288 clusters with a group of ST-474/CC-48 isolates because it was PCR-positive for a group of LOS locus class B2 orfs in addition to the LOS locus class E/H orfs typical of other ST/CC-45 isolates. Conversely, the ST-474/CC-48 isolates that H288 clusters with were PCR-positive for a group of LOS locus class E orfs in addition to the LOS locus class B2 orfs typical of other ST-474/CC-48 isolates. The second anomalous ST/CC-45 isolate, P115b, clusters with ST/CC-354 isolates because it is PCR-negative for all the ST/CC-45-characteristic LOS locus class E/H orfs, and PCR-positive for some ST/CC-354 characteristic orfs. Both anomalous ST/CC-354 isolates (T15 and P303a) that cluster with ST/CC-45 isolates are PCR-negative for all their ST/CC-354 characteristic orfs and PCR-positive for LOS locus class E-characteristic orfs.

Of the 46 CC-48 isolates we examined, 44 isolates were PCR-positive for a characteristic, extended set of ten LOS locus class B2 orfs (orf3bf, orf4ab, orf5ab2, orf6ab2, orf7ab–orf11ab, orf5bII), and were also PCR-positive for orf3c (Fig. 1, Supplementary Table S1). The two exceptions, human isolates H73020 and H164, were both PCR-negative for orf4ab and orf3c. While the majority (54%) of CC-48 isolates, and all non-ST-474 isolates within CC-48, were PCR-positive for only the ten LOS locus class B-characteristic orfs and orf3c, 14 of 35 ST-474 isolates were also PCR-positive for one or more LOS locus class E-characteristic orfs (orf3e–orf34e), and a partially overlapping set of 13 ST-474 isolates were PCR-positive for orf20df. Most prominent of these is a sub-cluster of five isolates, comprising four human isolates (H22082, H123, H128, H133), and a retail poultry isolate from company 2 (P238b), that were PCR-positive for a group of ten LOS locus class

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

E-characteristic orfs (orf3e–orf27e and orf29e–orf32e), but PCR-negative for orf28e. Two further retail poultry isolates from company 1, P110b and P170c, were PCR-positive for subgroups of six and eight LOS locus class E-characteristic orfs, respectively, but could be further distinguished from the five-isolate sub-cluster by being PCR-positive for orf28e.

ST/CC-45 isolates sourced from wild-bird faeces collected from recreational areas displayed a single binary genomotype (BG 23), while isolates from human clinical cases and poultry displayed five and three binary genomotypes respectively (Fig. 1, Supplementary Table S1). Of the three poultry binary genomotypes, the first (BG 26) was seen in isolates from companies 1 and 2 (P10b and P11b), the second (BG 20) in a single isolate from company 3 (isolate P115b), and the third (BG 22) in two isolates from each of companies 2 and 3. All ST/CC-45 isolates we examined from wild-bird faeces were PCR-positive for orf15c and orf26e, while all isolates we examined from retail poultry were PCR-negative for orf15c and orf26e. The binary genomotype of human isolate H051206734 (BG 23) was identical to that of the four isolates from wild-bird faeces (R3, R48D, R53B and R5a), suggesting that infection could have been acquired following contact with wild-bird faeces. In contrast, the binary genomotype of human isolate H73687 (BG 22) was identical to that of four poultry isolates (P133a, P135b, P137c and P155b), so was most likely acquired from a retail poultry source. Assignment of the most likely source was more equivocal for four additional human ST/CC-45 isolates, H75250a, H76800, H272 and H288. While isolates H75250a and H76800 cluster most closely in Figure 1 with ST/CC-45 poultry isolates, albeit as a discrete sub-cluster, and H288 (BG 5) clusters with ST-474 isolates, each of the four isolates was, in common with the wild-bird isolates, PCR-positive for orf26e. The presence or absence of orf26e also distinguishes between LOS locus classes E (orf26e present) and H (orf26e absent). Thus, all ST/CC-45 isolates from wild-bird faeces were assigned to LOS locus class E while, with the exception of isolate P115b (which could not be assigned), all ST/CC-45 isolates from retail poultry were assigned to LOS locus class H. Each of H051206734, H75250a, H76800 and H272 was assigned to LOS locus class E. The presence of LOS locus class E-characteristic orfs in ST/CC-45 isolate H288 was complicated by the presence of an additional group of LOS locus class B2 orfs.

ST-190/CC-21 isolates all display a core binary genomotype comprising orfs characteristic of LOS locus class C (orf3c-orf16c), but display significant mosaicism in gene complement on an isolate by isolate basis (Fig. 1, Supplementary Table S1). Two distinct clusters of ST-190/CC-21 isolates were observed. The smaller cluster, including three human isolates and a poultry isolate, is characterized by the common LOS locus class C orfs, together with orf19df–orf28e. The second, larger, eight-isolate cluster also includes both poultry and human isolates, and is characterized by the common LOS locus class C orfs, and only those orfs for strain H787, plus a small number of additional orfs, principally one or more of orf20df, orf3e and orf22e, which are in common with the smaller, four-isolate cluster.

All but two isolates we examined from ST/CC-354 were PCR-positive for a common core of five orfs (orf3bf, orf3c, orf18df, orf19df and orf20df), with some isolates displaying additional, limited mosaicism beyond the five-orf core (Fig. 1, Supplementary Table S1). However, two ST/CC-354 isolates (T15 and P303a) were PCR-negative for the five-orf core, and were instead PCR-positive for LOS locus class E orfs and, consequently, clustered with ST/CC-45 isolates.

DISCUSSION

In the current study we adopted a PCR-based approach to binary genomotyping, assessing LOS gene complement by scoring the full set of 42 well characterized LOS orfs described by Parker *et al.* [17] as either PCR-positive (orf present) or PCR-negative (orf absent or divergent). We also assigned LOS locus classes based upon a subgroup of 16 of the same orfs [17]. We applied this method to isolates from predominant poultry-associated STs (ST-474, ST-354, ST-45 and ST-190) identified in human clinical campylobacteriosis cases during the course of a sentinel surveillance study conducted from 2005 to 2008 in the Manawatu District of the lower North Island of New Zealand [22–24]. We determined that *C. jejuni* isolates indistinguishable on the basis of MLST can be distinguished on the basis of variation in their LOS orf complement and that these differences in orf complement can convey epidemiologically useful information.

We observed several instances of possible orf loss and/or gain. For example, while most ST/CC-354 isolates examined were PCR-positive for a common

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

core of five orfs (orf3bf, orf3c, orf18df, orf19df and orf 20df), two ST/CC-354 isolates (T15 and P303a) were PCR-negative for the five orfs, and were instead PCR-positive for LOS locus class E orfs and, consequently, clustered with ST/CC-45 isolates. This suggests that T15 and P303a have lost the core ST/CC-354 specific orfs and acquired a set of LOS locus class E orfs. We also observed several instances in which some isolates were PCR-positive for ST/CC characteristic orfs, but were also PCR-positive for additional orfs. For example, ST/CC-45 isolate H288 clusters with a group of ST-474/CC-48 isolates because it was PCR-positive for a group of LOS locus class B2 orfs in addition to the LOS locus class E/H orfs typical of other ST/CC-45 isolates. While this may indicate that H288 has acquired a set of LOS locus class B2 orfs resulting in a mosaic LOS locus, and although the DNA sample for H288 and other similar examples was prepared after a third round of single colony purification, we are not able to exclude the possibility that the *C. jejuni* colony was mixed, containing, in this example, both ST-474 and ST-45 strains. Nevertheless, Parker *et al.* [16] have previously described examples of mosaic LOS loci, so the current study may provide additional examples of such loci.

Most isolates we examined within CC-48 were PCR-positive for a characteristic, extended set of ten LOS locus class B2 orfs (orf3bf, orf4ab, orf5ab2, orf6ab2, orf7ab–orf11ab and orf5bII), and were also PCR-positive for orf3c. However, 40% of ST-474 isolates were also PCR-positive for one or more LOS locus class E orfs. In this latter group of ST-474 isolates were two particularly interesting sub-clusters classified in LOS locus class B2E that showed evidence for acquisition of a set of LOS locus class E orfs: Cluster 1 (BGs 3 and 4), comprising four human isolates and a retail poultry isolate from poultry company 2 that were PCR-positive for a group of ten LOS locus class E-characteristic orfs, but PCR-negative for orf28e; and Cluster 2 (BGs 1 and 2), comprising retail poultry isolates from company 1 that were PCR-positive for fewer LOS locus class E-characteristic orfs, but were PCR-positive for orf28e. Although occasionally isolated in retail poultry from company 2, ST-474 is strongly associated with retail poultry from company 1 [22, 24]. ST-474, along with other poultry-associated STs, was also associated with a winter campylobacteriosis epidemic in New Zealand in May, June, and to a lesser extent, July 2006 [2, 22], with three of the seven ST-474 LOS locus class B2E isolates being derived from that time period.

Isolates within each of ST/CC-45, ST/CC-354 and ST-190/CC-21 displayed intra-ST mosaicism in their LOS orf complement leading to epidemiologically useful binary genomotypes. Isolates we examined from ST/CC-45 in particular highlight the potential utility of LOS orf-based binary genomotyping as an extension to standard MLST. Isolates assigned to ST/CC-45 have been isolated from several sources in the Manawatu District, including retail poultry and wild-bird faeces from recreational areas, but source-attribution models based upon MLST allelic profiles and their prevalence in individual sources can only assign ST/CC-45 human clinical isolates to those sources on a probabilistic basis [22–27]. Here, binary genomotyping provided additional evidence for associating isolates from human clinical cases to specific sources, for example linking one human clinical isolate to poultry sources, and four human clinical isolates to wild-bird sources. French *et al.* [36] have recently shown that ST-45 was the most common single ST isolated from wild-bird faecal material in a study of children's playgrounds in Palmerston North, the principal city in the Manawatu District. Furthermore, among the three binary genomotypes we observed for ST/CC-45 poultry isolates, the first was seen in isolates from Companies 1 and 2 (P10b and P11b), the second in a single isolate from company 3 (isolate P115b), and the third in two isolates from each of companies 2 and 3, suggesting that, either transmission occurred between flocks maintained by these companies, or they acquired these strains from a common source.

We observed two distinct clusters of ST-190/CC-21 isolates, the first an almost identical cluster comprising three human isolates and a poultry isolate, and a second more variable cluster, also comprising human and poultry isolates (Fig. 1). Poultry isolates within both clusters were sourced from retail poultry meat from a single poultry company. We also observed binary genomotypes in ST-190/CC-21 isolates from human clinical cases that were not observed in poultry isolates. ST-190/CC-21 isolates have also been obtained from ruminants [22, 24], which may be an alternative source for some of the human isolates in the variable cluster. While we did examine a single ST-190/CC-21 isolate from retail beef (M1272a), this isolate clustered alongside a poultry isolate (P165a) that did not have an identical binary genomotype, and that was isolated from retail poultry at a different time

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

and from a different retail outlet to the meat sample. Nevertheless, clustering of these two isolates suggests that the two host species may sometimes acquire *Campylobacter* from a common environmental source.

While two isolates from ST/CC-354 were placed in LOS locus class E, and one isolate was placed in LOS locus class D, the remaining nine isolates (class U) we examined could not be placed within a formal LOS locus class. LOS locus classes D and F can be distinguished by the presence (D) or absence (F) of orf17d. With orf17d being PCR-negative in the ST/CC-354 class U isolates, those isolates likely represent variants of LOS locus class F. With additional LOS locus sequences being regularly added to the GenBank database as components of whole-genome sequences or as specific LOS locus sequences [16], it may be possible to design PCR primers to formally classify these isolates in the future.

While genome sequencing and partial genome sequencing of *C. jejuni* strains has shown that the LOS locus is hypervariable in gene complement when considered across all strains [10–19], our data, comparative genomic DNA microarray data [37], and the data from sequenced/partially sequenced strains when considered in its ST/CC context, show that intra-ST mosaicism in LOS gene complement observed on an isolate by isolate level is seen alongside broader patterns of congruence in gene complement that distinguish between different STs/CCs. These patterns of congruence in gene complement at the ST/CC level are reflected, in part, by similarities in assignment of LOS locus class within CC-48, ST/CC-45 and ST-190/CC-21. Indeed, consistent with our observations, a recent investigation of Belgian *C. jejuni* isolates assigned CC-21 isolates primarily to LOS locus class C, CC-45 isolates primarily to LOS locus class E, and CC-354 isolates primarily to LOS locus class D [8]. Assignment of LOS locus classes is also indicative of likely structural features of the LOS produced by an isolate. Thus, the sialylated LOS locus classes A, B and C are particularly associated with GBS and Miller–Fisher syndrome [17, 38, 39], and were assigned to isolates within CC-48 (class B2 and B2E) and ST-190/CC-21 (class C and CE) in the current study. However, we have no record that any of the human isolates we examined were derived from patients displaying symptoms of GBS or Miller–Fisher syndrome.

The current study demonstrates that binary genomotyping using PCR primer pairs that amplify a well characterized set of 42 LOS genes/orfs [17] serves as an epidemiologically useful extension to MLST that captures patterns of both mosaicism and congruence in gene complement. While this study focused on developing binary genomotyping as a method to subtype isolates from poultry-associated STs using a broad set of 42 LOS orfs, it is evident from our data that intra-ST subtyping could be performed using a subset of the 42 orfs we examined, eliminating orfs that are uninformative for a given ST, thereby significantly reducing the cost. Furthermore, while we have assigned numerical binary genomotypes to each unique 42-orf pattern we observed, PCR-based binary genomotyping also lends itself to recording data in a binary, digital format (1 for PCR positive, 0 for PCR negative). Such digital outputs create reproducible and portable binary genomotypes that can be readily formatted in databases, facilitating data storage and comparison between laboratories. Thus, PCR-based binary genomotyping lends itself to implementation in any PCR-capable laboratory, and offers a targeted glimpse at the underlying genetic structure of isolates being examined. Furthermore, because MLST is based upon sequencing regions from seven housekeeping genes, it is disconnected from effectors of isolate virulence. Bacteria often utilize variation in the structure of surface-exposed macromolecules as direct effectors of virulence or as a mechanism to evade the host immune system. Existing extensions to MLST involving sequencing of regions within more variable genes such as the *flaA* SVR, the *flaB* SVR and *porA* [29] already exploit such variation, and PCR-based binary genomotyping assessing variations in the complement of LOS biosynthesis genes provides an avenue for extension of MLST additional to those three approaches.

ACKNOWLEDGEMENTS

This work was supported through the Avian Health Fund and the Building Research Capabilities in Strategically Relevant Areas (BRCSRA) Fund. This paper is dedicated to the memory of our friend and colleague Dr Grant Hotter who died suddenly on 3 October 2009.

NOTE

Supplementary material accompanies this paper on the Journal's website (http://journals.cambridge.org/hyg).

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

DECLARATION OF INTEREST

None.

REFERENCES

1. **Anon.** Notifiable and other diseases in New Zealand. Annual Report 2005. ESR Ltd, Wellington New Zealand, 2006 (http://www.surv.esr.cri.nz/PDF_surveillance/AnnualRpt/AnnualSurv/2005AnnSurvRpt.pdf). Accessed 16 September 2009.
2. **Anon.** Notifiable and other diseases in New Zealand. Annual Report 2007. ESR Ltd, Wellington New Zealand, 2008 (http://www.surv.esr.cri.nz/PDF_surveillance/AnnualRpt/AnnualSurv/2007AnnualSurvRpt.pdf). Accessed 16 September 2009.
3. **Anon.** *Campylobacter* risk management strategy. New Zealand Food Safety Authority, Wellington, New Zealand, 2008 (http://www.nzfsa.govt.nz/foodborne-illness/campylobacter/strategy/Campylobacter_risk_management_strategy_2008-2011.pdf). Accessed 16 September 2009.
4. **Anon.** New Zealand public health surveillance report March 2009: covering October-December 2008. ESR Ltd, Wellington New Zealand, 2009 (http://www.surv.esr.cri.nz/PDF_surveillance/NZPHSR/2009/NZPHSR2009March.pdf). Accessed 16 September 2009.
5. **Yuki N.** Carbohydrate mimicry: a new paradigm of autoimmune diseases. *Current Opinion in Immunology* 2005; **17**: 577–582.
6. **Yuki N, *et al.*** Carbohydrate mimicry between human ganglioside GM1 and *Campylobacter jejuni* lipooligosaccharide causes Guillain–Barré syndrome. *Proceedings of the National Academy of Sciences USA* 2004; **101**: 11404–11409.
7. **Lake R, *et al.*** Lack of association between long-term illness and infectious intestinal disease in New Zealand. *New Zealand Medical Journal* 2004; **117** (http://www.nzma.org.nz/journal/117-1194/893/content.pdf).
8. **Habib I, *et al.*** Correlation between genotypic diversity, lipooligosaccharide gene locus class variation, and caco-2 cell invasion potential of *Campylobacter jejuni* isolates from chicken meat and humans: contribution to virulotyping. *Applied and Environmental Microbiology* 2009; **75**: 4277–4288.
9. **Louwen R, *et al.*** The sialylated lipooligosaccharide outer core in *Campylobacter jejuni* is and important determinant for epithelial cell invasion. *Infection and Immunity* 2008; **76**: 4431–4438.
10. **Champion OL, *et al.*** Comparative phylogenomics of the food-borne pathogen *Campylobacter jejuni* reveals genetic markers predictive of infection source. *Proceedings of the National Academy of Sciences USA* 2005; **102**: 16043–16048.
11. **Dorrell N, *et al.*** Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. *Genome Research* 2006; **11**: 1706–1715.
12. **Fouts DE, *et al.*** Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. *PLOS Biology* 2005; **3**: e15.
13. **Gilbert M, *et al.*** The genetic basis for the variation in the lipo-oligosaccharide of the mucosal pathogen, *Campylobacter jejuni*. *Journal of Biological Chemistry* 2002; **277**: 327–337.
14. **Hofreuter D, *et al.*** Unique features of a highly pathogenic *Campylobacter jejuni* strain. *Infection and Immunity* 2006; **74**: 4694–4707.
15. **Karlyshev A, *et al.*** Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Molecular Microbiology* 2005; **55**: 90–103.
16. **Parker CT, *et al.*** Characterization of lipooligosaccharide-biosynthetic loci of *Campylobacter jejuni* reveals new lipooligosaccharide classes: evidence of mosaic organizations. *Journal of Bacteriology* 2008; **190**: 5681–5689.
17. **Parker CT, *et al.*** Comparison of *Campylobacter jejuni* lipooligosaccharide biosynthesis loci from a variety of sources. *Journal of Clinical Microbiology* 2005; **43**: 2771–2781.
18. **Parkhill J, *et al.*** The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 2000; **403**: 665–668.
19. **Quiñones B, *et al.*** Comparative genomic analysis of clinical strains of *Campylobacter jejuni* from South Africa. *PLOS One* 2008; **3**: e2015.
20. **Dingle KE, *et al.*** Multilocus sequence typing system for *Campylobacter jejuni*. *Journal of Clinical Microbiology* 2001; **39**: 14–23.
21. **On SL, *et al.*** Molecular epidemiology of *Campylobacter* species. In: Nachamkin I, Szymanski CM, Blaser MJ, eds. *Campylobacter*. Washington, DC: ASM Press, 2008, pp. 191–211.
22. **French N.** Enhancing surveillance of potentially foodborne enteric diseases in New Zealand: human campylobacteriosis in the Manawatu. 2008 (http://www.nzfsa.govt.nz/science/research-projects/Campy_Attribution_Manawatu.pdf). Accessed 16 September 2009.
23. **Mullner P, *et al.*** Source attribution of food-borne zoonoses in New Zealand: a modified hald model. *Risk Analysis* 2009; **29**: 970–984.
24. **Mullner P, *et al.*** Assigning the source of human campylobacteriosis in new Zealand: acomparative genetic and epidemiological approach. *Infection, Genetics and Evolution* (in press).
25. **Sheppard SK, *et al.*** *Campylobacter* genotyping to determine the source of human infection. *Clinical Infectious Diseases* 2009; **48**: 1072–1078.
26. **Strachan NJC, *et al.*** Attribution of campylobacter infections in northeast Scotland to specific sources by use of multilocus sequence typing. *Journal of Infectious Diseases* 2009; **199**: 1205–1208.
27. **Wilson DJ, *et al.*** Tracing the Source of Campylobacteriosis. *PLoS Genetics* 2008; **4**: e1000203.
28. **McTavish SM, *et al.*** Wide geographical distribution of internationally rare *Campylobacter* clones within

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press

New Zealand. *Epidemiology and Infection* 2008; **136**: 1244–1252.

29. **Dingle KE, *et al.*** Extended sequence typing of *Campylobacter* spp., United Kingdom. *Emerging Infectious Disease* 2008; **14**: 1620–1622.
30. **Kärenlampi R, Rautelin H, Hänninen M-L.** Evaluation of genetic markers and molecular typing methods for prediction of sources of *Campylobacter jejuni* and *Campylobacter coli* infections. *Applied and Environmental Microbiology* 2007; **73**: 1683–1685.
31. **Merchant-Patel S, *et al.*** Characterisation of chicken *Campylobacter jejuni* isolates using resolution optimised single nucleotide polymorphisms and binary gene markers. *International Journal of Food Microbiology* 2008; **128**: 304–308.
32. **Price EP, Huygens F, Giffard PM.** Fingerprinting of *Campylobacter jejuni* by using resolution-optimised binary gene targets from comparative genomic hybridization studies. *Applied and Environmental Microbiology* 2006; **72**: 7793–7803.
33. **Dagerhamn J, *et al.*** Determination of accessory gene patterns predicts the same relatedness among strains of *Streptococcus pneumoniae* as sequencing of housekeeping genes does and represents a novel approach in molecular epidemiology. *Journal of Clinical Microbiology* 2008; **46**: 863–868.
34. **Laing C, *et al.*** Rapid determination of *Escherichia coli* O157:H7 lineage types and molecular subtypes by using comparative genomic fingerprinting. *Applied and Environmental Microbiology* 2008; **74**: 6606–6615.
35. **Sokal RR, Michener CD.** A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 1958; **38**: 1409–1438.
36. **French NP, *et al.*** Molecular epidemiology of *Campylobacter jejuni* isolates from wild-bird fecal material in children's playgrounds. *Applied and Environmental Microbiology* 2009; **75**: 779–783.
37. **Taboada EN, *et al.*** Comparative genomic assessment of multi-locus sequence typing: rapid accumulation of genomic heterogeneity among clonal isolates of *Campylobacter jejuni*. *BMC Evolutionary Biology* 2008; **8**: 229.
38. **Godschalk PCR, *et al.*** Structural characterization of *Campylobacter jejuni* lipooligosaccharide outer cores associated with Guillain-Barré and Miller–Fisher syndromes. *Infection and Immunity* 2007; **75**: 1245–1254.
39. **Godschalk PCR, *et al.*** PCR-restriction fragment length polymorphism analysis of *Campylobacter jejuni* genes involved in lipooligosaccharide biosynthesis identifies putative molecular markers for Guillain-Barré syndrome. *Journal of Clinical Microbiology* 2007; **45**: 2316–2320.

https://doi.org/10.1017/S0950268809991075 Published online by Cambridge University Press